Electronic Supplementary Material (ESI) for ChemComm.

Supporting Information

A bifunctional chemical signature enabling RNA 4-thiouridine enrichment sequencing with single-base resolution

Li Su,‡ Fengjiao Zhao,‡ Huahang Yu, Min Bai, Jing Xue, Xiaowen Cao, Feng Chen, and Yongxi Zhao*

Institute of Analytical Chemistry and Instrument for Life Science, The Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xianning West Road, Xi'an, Shaanxi 710049 (China). E-mail: yxzhao@mail.xjtu.edu.cn

‡ These authors contributed equally to this work.

Table of Contents

Experimental Sections

Materials

All commercially available materials were purchased from the indicated suppliers and used without further purification. 4-Thiouridine (4sU) was obtained from Aladdin Reagent Co. Ltd (Shanghai, China). 4-Thiouridine-5'-Triphosphate (4sUTP) was obtained from Trilink Biotechnologies (California, USA). N-Hex-5-ynyl-2-iodo-acetamide (E-IAA) was bought from J&K (Beijing, China). DNAse I, Taq DNA Polymerase and Mighty TA-cloning Kit were obtained from Takara Biotechnology Co. Ltd (Dalian, China). Shrimp Alkaline Phosphatase (rSAP), Bst 2.0 DNA polymerase, T7 RNA polymerase, PhotoScript II，Bst 3.0 DNA polymerase, Hot Start Taq DNA Polymerase and NotI-HF were purchased from New England Biolabs Ltd (Beijing, China). Dynabeads™ MyOne™ Streptavidin C1, SuperScript II, RNase I and Total RNA Extractor（Trizol） were obtained from ThermoFisher Scientific. GoScript Reverse Transcriptase was purchased from Promega. Disulphide Biotin Azide was purchased from Click Chemistry Tools (Scottsdale, USA). The oligonucleotides used in this work (listed in Table S1) were all synthesized and purified by Sangon Biotech Company Ltd. (Shanghai, China).

A bifunctional chemical signature for 4sU nucleoside and 4sU-RNA substrates with E-IAA

The bifunctional chemical signature of 4sU nucleoside was performed in 10% DMSO, 50 mM sodiumphosphate buffer pH 8, 20 mM E-IAA for 30 min at 50 °C according to a previous literature.[1] To obtain 4sU-RNA substrates, we used in vitro transcription of dsDNA templates (Table S1) catalyzed by T7 RNA polymerase. This reaction used 4sUTP in place of UTP, yielding ssRNA containing 4sU sites. The ssRNA products were purified after the degradation of dsDNA templates. For a typical reaction of 4sU-RNA substrates, 4sU-RNA was reacted with 20 mM E-IAA and 50 mM sodiumphosphate at pH 8 under 50 °C for 30 min. The reaction products were either degraded for LC-MS/MS analysis or subjected to reverse transcription for Sanger sequencing and restriction endonuclease assay (see below).

Mass spectrometry analysis

For the measurement of reaction products of 4sU nucleoside and E-IAA, the result was analyzed by High-resolution mass spectrum, HRMS (ESI) in the negative ion mode. For the measurement of 4sU-RNA reaction products, 8 μg 4sU-RNA or U-RNA was incubated with 20 mM E-IAA in 20 μl 50 mM sodiumphosphate buffer (PH=8) at 50 °C for 30 min in 10% DMSO. After the reaction, we used ethanol precipitation to purify the RNA to remove excess chemicals. The samples were digested into nucleosides by Shrimp Alkaline Phosphatase (rSAP) and RNase I through incubation for 12 h. After digestion, we carried aftertreatment according previously described.[2] 2 uL of the solution was injected into LC-MS/MS. the samples were analyzed by a Waters I-Class Vion IMS Qt. The nucleosides were separated by ultra-performance liquid chromatography on a reversephase column (1.8 um, 1.0×100 mm) in the negative ion mode. Liquid chromatography was performed by a gradient analysis: flow rate: 0.1 ml/min; eluent A: 0.01 TEAA in 3% (v/v) Acetonitrile; eluent B: Acetonitrile : H_2O=9:1(v/v); gradient: 0 min-4 min, 100% A, 0% B, curve: initial; 4 min-5 min, 90% A, 10% B , curve: 6; 5 min-12 min, 0% A, 100% B, curve: 6; 12 min-14 min, 100% A, 0% B, curve: 6.

Labeled 4sU was biotinylated by click reaction with disulphide biotin azide. The reactions were incubated at 37 °C in the dark for 1 h with 2.5 mM disulphide biotin azide, 1 mM $CuSO_4$, and 5 mM sodium ascorbate. After reaction, 100 mM DTT was added into the solution to break the disulphide bond and produce triazole nucleoside. The results were characterized by High-resolution mass spectra in negative ion mode.

Enrichment of single 4sU-RNA

Two fluorescently labeled RNAs were prepared for 4sU enrichment: containing single 4sU RNA was hybridized by CY5-4sU(5-Cy5-TGAGCCTTCCTCGGTACGGTCTGTA-3) through hybridization of Hyb-4sU (5-TCTTTGCTCCGTGTCCTGCCTACAGACCGTACCGAGGAAG GCTCA-3). Containing single U RNA was hybridized

by FAM-U (5-FAM- GCGTATAATAATGCTGGGCTCTAGT -3) through hybridization of Hyb-U (TCTTTGCTCCGTGTCCTGCCACTAGAGCCCAGCATTATTATACGC). Labeling 4sU-RNA was performed by E-IAA as described above. After alcohol precipitation, the labeled 4sU-RNA was biotinylated by click reaction with disulphide biotin azide. 2.5 mM disulphide biotin azide, 1 mM $CuSO_4$, 5 mM sodium ascorbate. Reactions were incubated at 37 °C in the dark for 1 h. After reaction, the samples were washed twice with 75% ethanol and 50% isopropanol respectively. Biotinylated RNA was separated from non-labeled RNA using Dynabeads MyOne Streptavidin C1 beads (Invitrogen). The operation of the experiment is based on the protocol of Dynabeads MyOne Streptavidin. Biotinylated RNA was incubated with 100 ul Dynabeads with rotation for 1 h at room temperature in the dark. Beads were magnetically fixed and washed twice with Dynabeads high-salt wash buffer (FT). 4sU-RNA was eluted with 100 ml of elution buffer (10 mM Tris [pH 7.4] and 100 mM DTT) (E). Fractions were concentrated and characterized by 15% native-PAGE.

4sU-RNA qPCR enrichment studies

Mixed samples with different proportions of 4sU-RNA at 1/10, 1/100, 1/1000, 1/10000 and 1/20000 were subjected to standard chemical reaction with E-IAA, then the RNA was purified with 75% ethanol and 50% isopropanol. The RNA was further performed for click chemistry to obtain biotinylation, the pull down procedure was used with Dynabeads MyOne Streptavidin C1 beads in a reported assay.[3] The operation of the experiment is based on the protocol. After that, magnetic beads were collected using a magnet (Promega) and the DTT eluant was Incubated at 37 °C for 30 min. DTT eluant was collected and dried by DNA Speedvac, then redissolved in 25 μL ddH_2O for further qPCR analysis. qPCRs were performed using a LightCycler 96-RealTime System. The enriched-RNAs were subjected to qPCR using protocol outlined by the manufacturer. RNA concentration was quantified by comparison with calibration lines of known concentration of input.

NotI restriction endonuclease assay

RNA containing a single 4sU site (one-4sU RNA) and multiple 4sU sites (multi-4sU RNA) were prepared by in vitro transcription as mentioned above. The reaction mixture was then treated with DNAse I at 37 °C for 1 h. The RNA products were purified by ethanol precipitation. 3.8 μg RNA was treated with 20 mM E-IAA under standard conditions. After ethanol precipitation, 50 ng RNA was subjected to reverse transcription reaction catalyzed by ProtoScript II Reverse Transcriptase, SuperScript II Reverse Transcriptase or Bst DNA polymerase. Then, 3 μL of the resulted solution was used for subsequent PCR amplification: 95 °C for 10 min, 30× (95 °C for 15 s, 56 °C for 20 s, 72 °C for 20 s), 72 °C for 5 min. The PCR products can be used for Sanger sequencing or restriction endonuclease assay. For the restriction endonuclease assay, the PCR products were first incubated with 10 U NotI in 1×Cutsmart buffer at 37 °C for 1 h. Then, the incubated mixture was detected by DNA polyacrylamide gel electrophoresis (PAGE) or DNA melting curve (LightCycler 96, Roche).

For biotinylated RNA, Labeled 4sU-RNA was biotinylated by click reaction with disulphide biotin azide, 2.5 mM disulphide biotin azide, 1 mM $CuSO_4$, and 5 mM sodium ascorbate. Reactions were incubated at 37 °C in the dark for 1 h. After reaction, 100 mM DTT was added into the solution to break the disulphide bond and produce triazolenucleoside. The samples were washed twice with 75% ethanol and 50% isopropanol respectively. Then, reverse transcription, PCR and enzyme digestion experiments as described above.

***E. coli* $tRNA^{Val}$ analysis**

Total RNA from E. coli DH5α cells (grown overnight in 3 ml standard Luria Bertani (LB) medium) were isolated using Total RNA Extractor (Trizol) according to the manufacturer's protocol. The total RNA was incubated with E-IAA as mentioned above. Reverse transcription of $tRNA^{Val}$, subcloning and sequencing were performed as previously described.[4] In brief, a stem loop-primer corresponding to the 5' end of $tRNA^{Val}$ was annealed to the cDNA of $tRNA^{Val}$ by the incubation at 65 °C for 5 minutes, followed by 10 minutes at 25 °C and immediate transferred to ice. The second strand of cDNA was synthesized by adding 5 mM dNTPs, 0.25 units Taq-Polymerase and Taq-buffer in a final volume of 10 μl and incubation at 25 °C, 60 °C and 72 °C for 2 minutes each. 5 μl of this reaction solution was used for subsequent PCR amplification. PCR

products were purified and subcloned. A total of 44 individual clones were randomly collected for the quantification of 4sU-to-C transition rate by Sanger sequencing of the plasmid DNA. All primers are listed in Table S1.

Table S1. Sequence information for oligonucleotides used in this study

Name	Sequences (**5′-3′**)
Sequence information for restriction-enzyme-digestion assay	
cDNA Sigle A	CTTGCGTTTCTCTCGTCCTTCTGTCTTGTGCTGTTGTTCGCGACCGCCCT GTGCGGTTG
cDNA Sigle G	CTTGCGTTTCTCTCGTCCTTCTGTCTTGTGCTGTTGTTCGCGGCCGCCCTGTGCGGTTG
cDNA T7	GAAATTAATACGACTCACTATAGGCAGGACACGGAGCAAAGACACAACCGCACAGG GCG
RT primer	CTTGCGTTTCTCTCGTCC
PCR fp1	TAGTGCTAGAGCATAGATGAGAAGACGAACGGCAGGACACGGAGCAAAGA
PCR rp1	AATGATAGCGCGACCACCGAGATCTACTTGCGTTTCTCTCGTCC
PCR fp2	AGGCCAAGAGTAGTCGAATTCCTCGTCCAGTAGTGCTAGAGCATAGATGA
PCR rp2	AGACCAAGCTACAACTACAGTCCGA AATGATAGCGCGACCACCGA
Sequence information for *E. coli* $tRNA^{Val}$	
$tRNA^{val}$ RT primer	GGTGGGTGATGACGGGATC
Stem-loop primer	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACGGGTGA
PCR fp	GTGCAGGGTCCGAGGT
PCR rp	GTGATGACGGGATC
Sequence information for enrichment of 4sU	
Hyb-U	TCTTTGCTCCGTGTCCTGCCACTAGAGCCCAGCATTATTATACGC
Hyb-4sU	TCTTTGCTCCGTGTCCTGCCTACAGACCGTACCGAGGAAGGCTCA
FAM-U	FAM- GCGTATAATAATGCTGGGCTCTAGT
CY5-4sU	Cy5-TGAGCCTTCCTCGGTACGGTCTGTA
Enrichment of 4sU	
q-pcr-U-cDNA mult-A	TGCCTTGCTCTACTCCTGTAATGTCTAGAGCTGTAGTACGCGACCGCCCTGTGCGGTT G
q-pcr-U-cDNAT7	GAAATTAATACGACTCACTATAGGACACCAGACGCAGCGCCAGGCAACCGCACAGG GCG
q-pcr-rt	TGCCTTGCTCTACTCCTG
q-pcr-fp1	AATGATAGCGCGACCACCGAGATCTATGCCTTGCTCTACTCCTG
q-pcr-rp1	TAGTGCTAGAGCATAGATGAGAAGACGAACGGACACCAGACGCAGCGCCA
Tm-A	TTGTTCGCGACCGCCCTG
Tm-G	TTGTTCGCGGCCGCCCTG

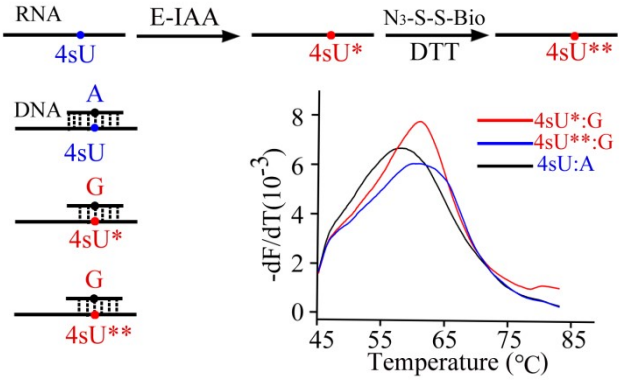

Figure S1. High-resolution melting (HRM) analysis of stability and selectivity of the base pairs about 4sU*:G and 4sU**:G. No obvious difference has been observed between 4sU*:G and 4sU**:G base pairs.

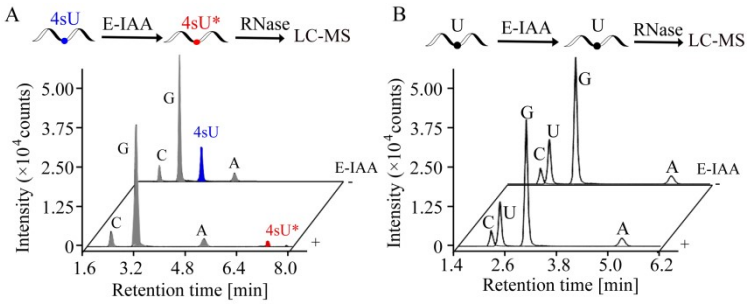

Figure S2. Analysis of products from the reaction of nucleosides and E-IAA. (A) LC-MS extracted ion chromatograms of nucleosides from treated 4sU-RNA and untreated 4sU-RNA with E-IAA. The reaction of 4sU and E-IAA may produce unknown by-products and we cannot extract molecular ion peak of unknown compounds. (B) LC-MS extracted ion chromatograms of nucleosides from treated U-RNA with E-IAA and untreated U-RNA.

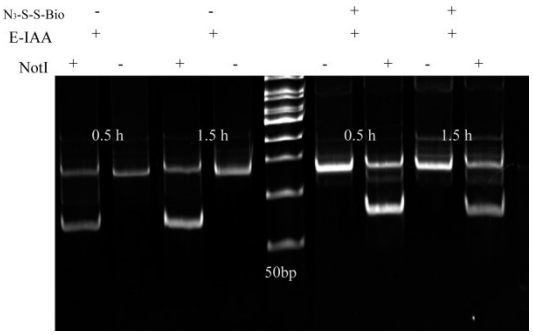

Figure S3. 15% native PAGE analysis of conversion efficiency of T-to-C with different chemical reagents reacting for 1.5h or 0.5h, respectively.

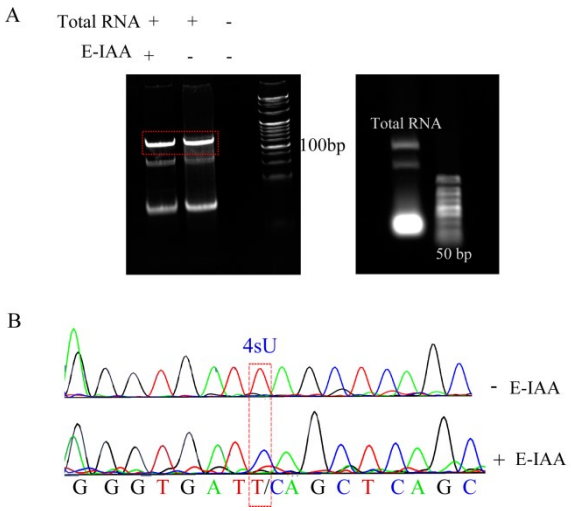

Figure S4. The 4sU-to-C transition performance of bacterial $tRNA^{Val}$. (A) Native PAGE analysis of RT/PCR products from $tRNA^{Val}$ isolated from *E. coli*. (B) Sanger sequencing of the PCR products. Notably, the sequence detected here is complementary to that detected in Figure 2D.

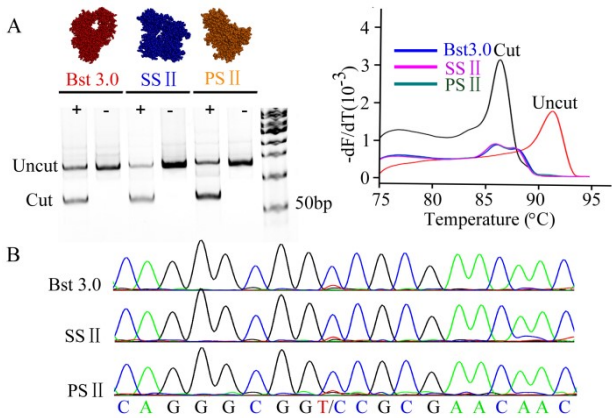

Figure S5. Detection of the 4sU at single-base resolution with different polymerase during reverse transcriptase processes. Native PAGE and melting curve (A) and Sanger sequencing (B) analysis of T-to-C conversion.

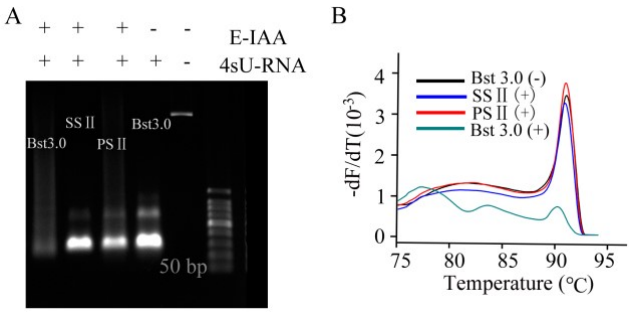

Figure S6. The effect of different enzymes used in RT on the primer extension for multiple 4sU-RNA. Agarose gel (A) and melting curve (B) analysis of polymerase action on labelled 4sU for multiple 4sU-RNA with different polymerases during reverse transcription.

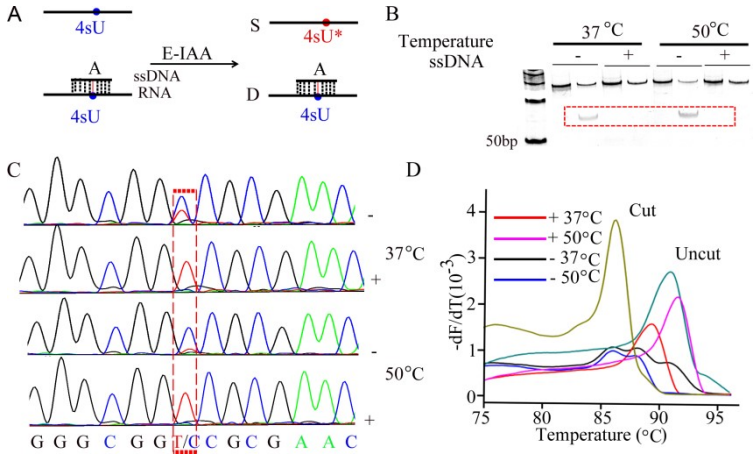

Figure S7. The inhibition effect of DNA/RNA duplex structure on the T-to-C transition. (A) Schematic illustration of the 4sU-to-C transition using RNA substrate and DNA/RNA duplex substrate. The DNA/RNA duplex is formed by the hybridization of RNA and ssDNA blocker. The 4sU site of RNA is in the complementary region. The reaction products were analyzed by 15% native PAGE (B), Sanger sequencing (C) and melting curve (D). These results indicate the RNA/DNA duplex structure significantly impeded the 4sU-to-C transition compared with RNA. It implied that the 4sU:A base pairing in RNA/DNA duplex prevented nucleophilic substitution of 4sU with E-IAA.

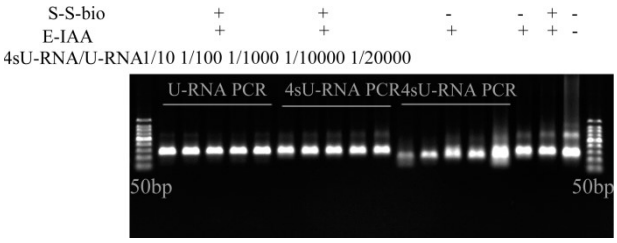

Figure S8. Agarose gel analysis of amplification products of PCR. 4sU-RNA/U-RNA mixed samples with different 4sU-RNA proportion of 1/10, 1/100, 1/1000, 1/10000 and 1/20000 were treated with E-IAA or E-IAA and disulphide biotin azide. Purified RNA was performed for RT/PCR reaction. Notably, integral DNA band can be detected from mixed samples for 4sU-RNA or U-RNA.

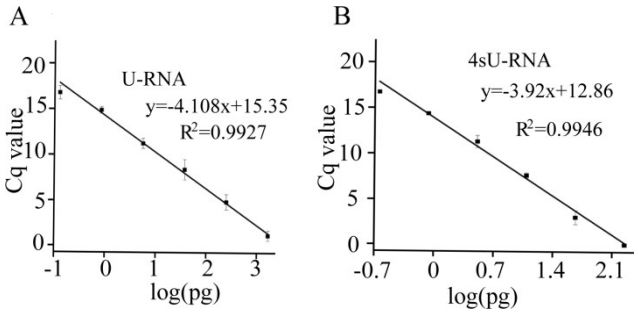

Figure S9. Example calibration line for U-RNA and 4sU-RNA. (A) Example calibration line for U-RNA. (B) Example calibration line for 4sU-RNA.

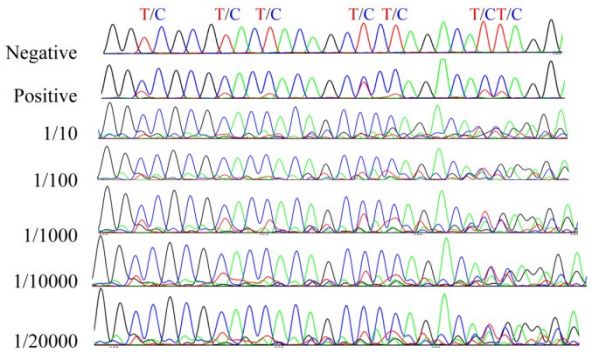

Figure S10. Sanger sequencing analysis of PCR products from enriched 4sU-RNA for mixed samples with different proportions of 4sU-RNA/U-RNA. 4sU-RNAs were amplified and then analysed by Sanger sequencing after enrichment. The 4sU site in RNA were detected at single base-resolution with enrichment.

Reference

1. V. A. Herzog, B. Reichholf, T. Neumann, P. Rescheneder, P. Bhat, T. R. Burkard, W. Wlotzka, A. von Haeseler, J. Zuber and S. L. Ameres, *Nat. Methods*, 2017, **14**, 1198-1204.
2. J. Spitzer, M.Hafner, M. Landthaler, M. Ascano, T. Farazi, G. Wardle, J. Nusbaum, M. Khorshid, L. Burger, M. Zavolan, T. Tuschl. *Methods Enzymol*, 2014, **539**, 113-161.
3. E. E. Duffy, M. Rutenberg-Schoenberg, C. D. Stark, R. R. Kitchen, M. B. Gerstein and M. D. Simon, *Mol. Cell*, 2015, **59**, 858-866.
4. C. Riml, T. Amort, D. Rieder, C. Gasser, A. Lusser and R. Micura, *Angew. Chem., Int. Ed*., 2017, **56**, 13479-13483.